

RD-R189 488

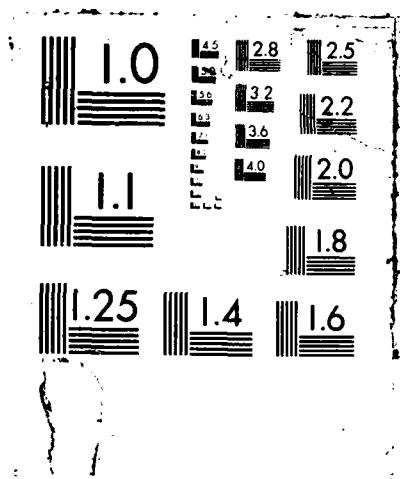
STOCHASTIC COMPARISONS OF ORDER STATISTICS WITH
APPLICATIONS IN RELIABILI. (U) FLORIDA STATE UNIV
TALLAHASSEE DEPT OF STATISTICS J S KIM ET AL. NOV 87
UNCLASSIFIED FSU-TR-M773 AFOSR-TR-87-217

1/1

F/G 12/3

NL





AD-A189 408

ARO 23699.16-MA

DTIC FILE COPY

(2)

STOCHASTIC COMPARISONS OF ORDER STATISTICS,
WITH APPLICATIONS IN RELIABILITY

by

Jee Soo Kim, Frank Proschan¹ and Jayaram Sethuraman²

GTE Laboratories
40 Sylvan Road
Waltham, MA 02154

and

Department of Statistics
Florida State University
Tallahassee, Florida 32306

and

Department of Statistics
Florida State University
Tallahassee, FL 30332

FSU Technical Report No. M-773
AFOSR Technical Report No. 87-217
USARO Technical Report No. D-99

November, 1987

DTIC
SELECTED
FEB 02 1988
S ~ E D

¹ Research supported by the Air Force Office of Scientific Research under Grant No. F49620-85-C-0007.

² Research supported by the Army Research Office under Grant DAAL03-86-K-0094.

Key Words and Phrases : Stochastic comparison, order statistics, reliability, life testing, majorization, Schur function, heterogeneous distributions, k -out-of- n system, proportional hazard functions.

88 1 27 1 34

**STOCHASTIC COMPARISONS OF ORDER STATISTICS,
WITH APPLICATIONS IN RELIABILITY**

by

Jee Soo Kim, Frank Proschan, and Jayaram Sethuraman

ABSTRACT

This is an invited paper to appear in the special issue of Communications in Statistics : Theory and Methods devoted to "Order Statistics and Applications". This paper reviews recent developments in stochastic comparisons of order statistics.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/ _____	
Availability Codes	
Dist	Avail and/or Special
A-1	



1. INTRODUCTION AND SUMMARY.

An impressive array of papers has been devoted to inequalities and stochastic ordering of linear combinations and partial sums of order statistics and comparison of their expectations. In this paper we survey recent advances in stochastic comparisons of order statistics along with reliability applications.

Section 2 presents inequalities for linear combinations of order statistics from restricted families obtained by Barlow and Proschan (1966a). Comparisons of linear combinations of order statistics from distributions F and G are obtained for $G^{-1}F$ convex and for $G^{-1}F$ starshaped. These results yield conservative upper and lower tolerance limits. For G exponential and F IFR or IFRA, we present stochastic comparisons for the "total time on test", used in life testing.

We introduce the notions of majorization and Schur function. Because majorization leads to many inequalities, these notions will be exploited extensively in the ensuing sections.

Section 3 presents stochastic comparisons of order statistics from underlying heterogeneous distributions. Given two sets of independent components (possibly unlike), majorization conditions are given by Pledger and Proschan (1971) which insure that any k -out-of- n system constructed from components in the first set will have reliability at least as great as that of a corresponding system constructed from components in the second set. Since the ordered failure times of the components represent order statistics from heterogeneous distributions, the order statistics from one set of underlying distributions $\{F_1, \dots, F_n\}$ can be compared stochastically with those from another set $\{F_1^*, \dots, F_n^*\}$.

We present additional comparisons involving spacings between order statistics. In some of the comparisons, the underlying heterogeneous distributions are compared with a single underlying homogeneous distribution, while in others, they are compared with another set of distributions less heterogeneous in the sense of majorization. These results of Pledger-Proschan can be used to approximate the reliability of certain types of systems of unlike components by computing the reliability of corresponding systems of like components.

The main theme of Section 4 is a result of Proschan and Sethuraman (1976). They stochastically compare whole vectors of order statistics, assuming F_1, \dots, F_n (F_1^*, \dots, F_n^*) have propor-

tional hazard functions with $\lambda_1, \dots, \lambda_n$ ($\lambda_1^*, \dots, \lambda_n^*$) as the constants of proportionality. Pledger and Proschan (1971) stochastically compare $X_{(r)}$ and $X_{(r)}^*$ separately for each r ; these results can be obtained as special cases of the Proschan-Sethuraman (1976) results.

Finally, an extension of stochastic comparison of random vectors to stochastic comparison of random processes is presented.

Throughout we use the term increasing (decreasing) for nondecreasing (nonincreasing).

2. INEQUALITIES FOR LINEAR COMBINATIONS OF ORDER STATISTICS.

Stochastic comparisons are made for linear combinations of order statistics from F and G when F is convex with respect to G (i.e., $G^{-1}F(x)$ is a convex function on the support of F , assumed an interval) and when F is starshaped with respect to G (i.e., $G^{-1}F(x)$ is a starshaped function). The concept of F being convex with respect to G was introduced in van Zwet (1964) and the concept of F being starshaped with respect to G is discussed in Barlow and Proschan (1981).

We adopt the following notation and assumption. Let $X(Y)$ have distribution $F(G)$. We assume that $F(0) = 0 = G(0)$, and that F and G are continuous. We assume also that the support of F is an interval, possibly infinite, and that G is strictly increasing on its support. We use \bar{F} for $1 - F$ and \bar{G} for $1 - G$.

A positive function h is starshaped on $[0, b]$, $0 < b \leq \infty$, if $h(x)/x$ is increasing for x in $[0, b]$, or equivalently, if $h(\alpha x) \leq \alpha h(x)$ for $0 \leq \alpha \leq 1, 0 \leq x < b$.

A failure rate $r(t)$ at time t is defined as $r(t) = f(t)/\bar{F}(t)$ when density $f(t)$ exists and $\bar{F}(t) > 0$. We say F has an increasing failure rate (IFR) if $r(t)$ is increasing in t and F has a decreasing failure rate (DFR) if $r(t)$ is decreasing in t .

A random variable X is said to be stochastically smaller than a random variable Y (denoted by $X \leq^{st} Y$) if $P(X > t) \leq P(Y > t)$ for every real number t . We say X is stochastically equal to Y ($X =^{st} Y$) if $P(x > t) = P(Y > t)$ for each real t . We stochastically compare linear combinations of order statistics $X_{(1)} \leq \dots \leq X_{(n)}$ from F and $Y_{(1)} \leq \dots \leq Y_{(n)}$ from G when $G^{-1}F$ is starshaped as well as when $G^{-1}F$ is convex on the support of F .

First we consider pairs of distributions F and G such that $G^{-1}F$ is starshaped on the support of F . Barlow and Proschan (1966a) proved the following two lemmas which are fundamental tools in obtaining stochastic inequalities. We shall find it convenient to define $\bar{A}_i = \sum_{j=i}^n a_j$, where the a_j represent real numbers.

Lemma 2.1. $h(\sum_{i=1}^n a_i x_i) \leq \sum_{i=1}^n a_i h(x_1)$ for all starshaped h on $[0, b)$ and all $0 \leq x_1 \leq \dots \leq x_n < b$

for which $0 \leq \sum_{i=1}^n a_i x_i < b$ if and only if there exists $k(1 \leq k \leq n)$ such that $0 \leq \bar{A}_1 \leq \dots \leq \bar{A}_k \leq 1$, and when $k < n$, $\bar{A}_{k+1} = \dots = \bar{A}_n = 0$.

Lemma 2.2. $h(\sum_{i=1}^n a_i x_i) \geq \sum_{i=1}^n a_i h(x_1)$ for all starshaped h on $[0, b)$ and all $0 \leq x_1 \leq \dots \leq x_n < b$

for which $0 \leq \sum_{i=1}^n a_i x_i < b$ if and only if there exists $k(1 \leq k \leq n)$ such that $\bar{A}_1 \geq \dots \geq \bar{A}_k \geq 1$; $\bar{A}_{k+1} = \dots = \bar{A}_n = 0$. If F is starshaped with respect to G , then by Lemma 2.1

$$G^{-1}F\left(\sum_{i=1}^n a_i X_{(i)}\right) \leq \sum_{i=1}^n a_i G^{-1}F(X_{(i)}) =^{st} \sum_{i=1}^n a_i Y_{(i)}.$$

This will be formally stated as follows.

Theorem 2.1. (Barlow and Proschan, 1966a). Let $G^{-1}F$ be starshaped on the support of F , $F(0) = 0 = G(0)$. If there exists $k(1 \leq k \leq n)$ such that $0 \leq \bar{A}_1 \leq \dots \leq \bar{A}_k \leq 1$, and when $k < n$, $\bar{A}_{k+1} = \dots = \bar{A}_n = 0$, then

$$F\left(\sum_{i=1}^n a_i X_{(i)}\right) \leq^{st} G\left(\sum_{i=1}^n a_i Y_{(i)}\right). \quad (2.1)$$

From Lemma 2.2 one may obtain the reverse inequality of (2.1). By assumption, the support of F is an interval, say $[0, b]$. If $\sum_{i=1}^n a_i X_{(i)} > b$, then $F(\sum_{i=1}^n a_i X_{(i)}) = 1 \geq G(\sum_{i=1}^n a_i Y_{(i)})$. Considering outcomes for which $\sum a_i X_{(i)} < b$, Lemma 2.2 leads to

$$G^{-1}F\left(\sum_{i=1}^n a_i X_{(i)}\right) \geq \sum_{i=1}^n a_i G^{-1}F(X_{(i)}) =^{st} \sum_{i=1}^n a_i Y_{(i)}.$$

The above discussion is summarized in the following theorem due to Barlow and Proschan.

Theorem 2.2. Let $G^{-1}F$ be starshaped on the support of F and $F(0) = 0 = G(0)$. Let $a_i \geq 0$ for $i = 1, 2, \dots, n$ and $a_n \geq 1$. Then

$$F\left(\sum_{i=1}^n a_i X_{(i)}\right) \geq^{st} G\left(\sum_{i=1}^n a_i Y_{(i)}\right). \quad (2.2)$$

Similar inequalities can be obtained for pairs of distributions F and G such that $G^{-1}F$ is convex on the support of F . This is a strengthening of the starshapedness hypothesis. Barlow and Proschan (1966a) present detailed discussions of the following theorems, which assume F is convex with respect to G .

Theorem 2.3. Let $G^{-1}F$ be convex on the support of F , $F(0) = 0 = G(0)$, and $0 \leq \bar{A}_i \leq 1$ for $i = 1, 2, \dots, n$. Then

$$F\left(\sum_{i=1}^n a_i X_{(i)}\right) \leq^{st} G\left(\sum_{i=1}^n a_i Y_{(i)}\right). \quad (2.3)$$

Theorem 2.4. Let $G^{-1}F$ be convex on the support of F , $F(0) = 0 = G(0)$, and for some $k (0 \leq k \leq n)$, $\bar{A}_i \geq 1$ for $i = 1, \dots, k$ and $\bar{A}_i \leq 0$ for $i = k+1, \dots, n$. Then

$$F\left(\sum_{i=1}^n a_i X_{(i)}\right) \geq^{st} G\left(\sum_{i=1}^n a_i Y_{(i)}\right). \quad (2.4)$$

From the simple identity $\sum_{i=1}^n a_i x_i = \sum_{i=1}^n \bar{a}_i (x_i - X_{(i-1)})$, one may verify (2.3) and (2.4) are equivalent to

$$F\left[\sum_{i=1}^n \bar{a}_i (X_{(i)} - X_{(i-1)})\right] \leq^{st} (\geq^{st}) G\left[\sum_{i=1}^n \bar{a}_i (Y_{(i)} - Y_{(i-1)})\right].$$

An important problem in statistical reliability theory and life testing is to obtain tolerance limits as a function of sample data. The above inequalities can be used to construct conservative lower or upper tolerance limits for IFR and IFRA distributions. Confidence limits for DFR and DFRA distributions can also be obtained using the same techniques. See Barlow and Proschan (1966b).

Let $G(t) = 1 - e^{-t}, t \geq 0$. Then $G^{-1}F$ starshaped on the support of F is equivalent to F having an increasing failure rate average (IFRA). We now discuss results concerning "total time on test" when successive observations are taken from an IRFA (DFRA) distribution.

If n items are put on life test and the test terminated at the time of the r^{th} failure (Type-II censored sampling), then $T_{rn} = \sum_{i=1}^n (n-i+1)(X_{(i)} - X_{(i-1)})$ represents the total time on test. This statistic has been extensively studied and applied in the case of the exponential distribution by Epstein and Sobel (1953) and Epstein (1960a,b). The best estimate of the mean θ in the exponential case is $\hat{\theta}(\underline{x}) = T_{rn}/r$.

Now let F be IFRA (DFRA), $F(0) = 0$, and $E(X) = \theta$. Then Barlow and Proschan (1966a) show that the total time on test divided by the sample mean associated with F is stochastically

larger (smaller) than that associated with the exponential distribution, i.e.,

$$\sum_{i=1}^r (n-i+1)(X_{(i)} - X_{(i-1)})/r\bar{X} \geq^{st} (\leq^{st}) \sum_{i=1}^r (n-i+1)(Y_{(i)} - Y_{(i-1)})/r\bar{Y}.$$

In the case of the exponential distribution, the normalized spacings $(n-i+1)(Y_{(i)} - Y_{(i-1)})$ are independent and identically distributed for $i = 1, \dots, n$ and $n \geq 1$. Thus one might expect that the spacings would exhibit certain monotonicity properties for distribution F such that $G^{-1}F$ is convex. Barlow and Proschan (1966a) show that if F is IFR (DFR) with $F(0) = 0$, then $(n-i+1)(X_{(i)} - X_{(i-1)})$ is stochastically increasing (decreasing) in $n \geq i$ for fixed i . We can establish as a corollary that if F is IFR(DFR), then the spacing $(n-i+1)(X_{(i)} - X_{(i-1)})$ is stochastically decreasing (increasing) in $i = 1, 2, \dots, n$ for fixed n . For example, $(n-i)(X_{(2)} - X_{(1)}) \leq^{st} nX_{(1)}$, where $X_{(0)} \equiv 0$. See Barlow, Marshall, and Proschan (1969) for further discussion of inequalities involving starshaped and convex functions. Chan, Proschan and Sethuraman (1983) consider an ordering different from the van Zwet convex ordering. They say that F is more convex than G , in symbols $F >^c G$, if $FG^{-1}(t)$ is a convex function on $(0,1)$. When F and G have density functions f and g , respectively, they show that $F >^c G$ if and only if $\frac{f(x)}{g(x)}$ is an increasing function of x , thus relating the concept of increasing likelihood ratios to convex ordering.

An extremely powerful and useful concept for deriving a great variety of inequalities is the concept of majorization.

Definition 2.1. Let $a_1 \leq a_2 \cdots a_n, b_1 \geq b_2 \geq \cdots b_n, \sum_{j=1}^k a_j \geq \sum_{j=1}^k b_j$ for $k = 1, 2, \dots, n-1$, and $\sum_{j=1}^n a_j = \sum_{j=1}^n b_j$. Then $\underline{a} = (a_1, \dots, a_n)$ is said to majorize $\underline{b} = (b_1, \dots, b_n)$, written $\underline{a} \geq^m \underline{b}$.

Definition 2.2. Let $\phi(\underline{a}) \geq \phi(\underline{b})$ whenever $\underline{a} \geq^m \underline{b}$. Then ϕ is called a Schur function.

Theorem 2.4. (Schur (1923), Ostrowski (1952)). Let $\phi(\underline{x})$ be a differentiable, real-valued function of n real variables. Then ϕ is a Schur function if and only if $(x_i - x_j)(\frac{\partial \phi}{\partial x_i} - \frac{\partial \phi}{\partial x_j}) \geq 0$ for all \underline{x} and all $i \neq j$.

For a systematic treatment of the theory of majorization and its applications in mathematics and statistics, see Marshall and Olkin (1979). The concept of majorization has been extended to elements of $L_1(0,1)$ in Ryff (1963) and the Schur-Ostrosowski theorem for Schur functionals has been obtained in Chan, Proschan and Sethuraman (1987).

Marshall, Olkin and Proschan (1967) determine conditions on (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) for the monotonicity of the ratio of means; $g(r) = (\sum_{i=1}^n a_i^r / \sum_{i=1}^n b_i^r)^{1/r}$. We next show that one application of this monotonicity yields a stochastic comparison between a function of order statistics from an IFRA distribution and the same function of order statistics from the exponential distribution. For notational simplicity, we write $(\frac{a_i^r}{\sum a_i^r}) = (\frac{a_1^r}{\sum a_i^r}, \dots, \frac{a_n^r}{\sum a_i^r})$.

We know that $\sum x \log x$ is a Schur function. Thus to show $g(r)$ is increasing in r , or equivalently, $d \log g(r)/dr \geq 0$, it is sufficient that $(\frac{\alpha}{\sum \alpha}) \geq^m (\frac{\beta}{\sum \beta})$, where $\alpha_i = a_i^r$ and $\beta_i = b_i^r$ (for $r \geq 0$) and where $\alpha_i = a_{n-i+1}^r$ and $\beta_i = b_{n-i+1}^r$ (for $r < 0$), $i = 1, 2, \dots, n$.

Marshall et al. (1967) obtain the following theorem.

Theorem 2.5. If $\alpha_1 > 0, \dots, \alpha_n > 0, \beta_1 \geq \dots \geq \beta_n > 0, \frac{\beta_1}{\alpha_1} \leq \dots \leq \frac{\beta_n}{\alpha_n}$, then $(\frac{\alpha}{\sum \alpha}) \geq^m (\frac{\beta}{\sum \beta})$. Using Theorems 2.4 and 2.5, Marshall et al. (1967) prove the monotonicity of $g(r)$. We will see how these results are used in making stochastic comparisons involving the order statistics from distributions F and G , where F is starshaped with respect to G .

An important example in which the conditions of Theorem 2.5 are satisfied is obtained by choosing $\alpha_i = \phi(\beta_i)$, where ϕ is a nonnegative starshaped function. We note that a nonnegative starshaped function ϕ must be increasing and must satisfy $\phi(0) = 0$. Such functions are discussed by Bruckner and Ostrow (1962). Assume that $\beta_1 \geq \dots \geq \beta_n > 0$; it follows that $\alpha_1 = \phi(\beta_1) \geq \dots \geq \alpha_n = \phi(\beta_n) \geq 0$. Thus by Theorem 2.5,

$$(\frac{\alpha}{\sum \alpha}) = (\frac{\phi(\beta)}{\sum \phi(\beta)}) \geq^m (\frac{\beta}{\sum \beta}). \quad (2.5)$$

Let $X_{(1)} \geq \dots \geq X_n$ be order statistics from F . Then from (2.5) we have for any starshaped function $\phi \geq 0$ that

$$(\frac{X_{(1)}}{\sum X_i}, \dots, \frac{X_n}{\sum X_i}) \leq^m (\frac{\phi(X_{(1)})}{\sum \phi(X_i)}, \dots, \frac{\phi(X_n)}{\sum \phi(X_i)}).$$

Consequently, if F is starshaped with respect to the distribution G , and $Y_1 \geq \dots \geq Y_n$ are order statistics from G , then $(\frac{X_{(1)}}{\sum X_i}, \dots, \frac{X_n}{\sum X_i}) \leq^s ((\frac{Y_1}{\sum Y_i}, \dots, \frac{Y_n}{\sum Y_i}))$, i.e., $\sum_{i=1}^k X_i / \sum_{i=1}^n X_i \leq^s \sum_{j=1}^k Y_j / \sum_{i=1}^n Y_i$ for $k = 1, 2, \dots, n$.

Theorem 2.4 yields some interesting applications. Choosing the Schur function $\phi(t_1, \dots, t_n) = n \sum t_i^2 - 1$, we obtain

$$\sum (X_i - \bar{X})^2 / \bar{X}^2 \leq^s \sum (Y_i - \bar{Y})^2 / \bar{Y}^2. \quad (2.6)$$

Choosing the Schur function $\phi(t_1, \dots, t_n) = n \sum A_i t_i / \sum t_i$, $a_1 \geq \dots \geq a_n, t_1 \geq \dots \geq t_n$, we obtain

$$\sum A_i X_i / \bar{X} \leq^{st} \sum A_i Y_i / \bar{Y} \quad (2.7)$$

An important special case where F is starshaped with respect to G is obtained by choosing $\bar{G}(x) = e^{-x}$ and F to be IFRA. The statistical applications of (2.6) in the problem of testing the hypothesis that F is exponential versus the alternative that F is IFRA, i.e., $\log \bar{F}(x)$ is concave where finite, and applications of (2.7) in testing for outliers when the distribution is known to be IFRA are discussed by Marshall et al (1967).

3. STOCHASTIC COMPARISONS OF ORDER STATISTICS FROM HETEROGENEOUS DISTRIBUTIONS.

A great body of statistical literature exists for order statistics from a single underlying distribution. See for example Sarhan and Greenberg (1962), Pyke (1965, 1970), David (1970, 1986), Groeneveld (1982), and references contained therein. The results involving order statistics from underlying heterogeneous distributions are far fewer. One motivation for considering underlying heterogeneous distributions arises in reliability theory, when one studies k -out-of- n systems. A system of n components is called a k -out-of- n system if it functions if and only if at least k components function. See Barlow and Proschan (1981). Note that the time of failure of a k -out-of- n system of independent components with respective life distributions F_1, F_2, \dots, F_n corresponds to the $(n - k + 1)^{th}$ order statistic from the set of underlying heterogeneous distributions $\{F_1, F_2, \dots, F_n\}$. Sen (1970) proved that the smallest (largest) order statistic of a sample of size n from heterogeneous populations is stochastically smaller (larger) than the smallest (largest) order statistic of a sample of size n from a common population whose distribution is the equally weighted mixture of the original distributions F_1, F_2, \dots, F_n . Let $X_{(1)} \leq \dots \leq X_n (Y_{(1)} \leq \dots \leq Y_{(n)})$ be the order statistics of n independent random variables $X_1, X_2, \dots, X_n (Y_1, Y_2, \dots, Y_n)$ with distribution functions $F_1, F_2, \dots, F_n (G_1, G_2, \dots, G_n)$ respectively. A result of Sen (1970) concerning stochastic relationship between order statistics is the following.

Theorem 3.1. Let $G_1(x) = \dots = G_n(x) = \frac{1}{n} \sum_{i=1}^n F_i(x)$. Then

$$X_{(1)} \leq^{st} Y_{(1)} \text{ and } X_{(n)} \geq^{st} Y_{(n)}.$$

Additional results are obtained by Pledger and Proschan (1971), wherein they assume that the distribution functions in the heterogeneous case have proportional hazard functions. In some of

their comparisons, the underlying heterogeneous distributions, less heterogeneous in the sense of majorization. Pledger and Proschan also compare stochastically spacing between order statistics.

The simplest comparisons of k -out-of- n systems can be made by taking fixed component reliabilities, p_1, p_2, \dots, p_n , rather than time-dependent component reliabilities $\bar{F}_1(t), \dots, \bar{F}_n(t)$. We denote the system reliability of a k -out-of- n system $h_k(p_1, \dots, p_n)$ as a function of component reliabilities p_1, \dots, p_n . For component reliability p_i we define the corresponding component hazard R_i by

$$R_i = -\log p_i. \quad (3.1)$$

By using notions of majorization and Schur function (see Section 2), Pledger and Proschan (1971) obtain the following inequalities.

Theorem 3.2. Let $\underline{R} = (R_1, \dots, R_n)$ be a vector of component hazards which majorizes $\underline{R}^* = (R_1^*, \dots, R_n^*)$, a second vector of component hazards. Then the corresponding reliabilities for a k -out-of- n system satisfy

$$h_k(\underline{p}) \geq h_k(\underline{p}^*) \text{ for } k = 1, \dots, n-1 \quad (3.2)$$

and

$$h_n(\underline{p}) = h_n(\underline{p}^*). \quad (3.3)$$

For a fixed vector \underline{p} of component reliabilities they also prove $h_k(\underline{p})$ is a Pólya frequency sequence of order $2(PF_2)$ in the index k , i.e., $h_k^2(\underline{p}) \geq h_{k-1}(\underline{p})h_{k+1}(\underline{p})$ for $k = 2, \dots, n-1$. See Karlin (1968) for discussion of PF_2 .

The results concerning time-dependent models follow immediately by setting $p_i = \bar{F}_i(t)$ and $p_i^* = \bar{F}_i^*(t)$. We assume independent observations, one observation from distribution $F_i(F_i^*), i = 1, 2, \dots, n$. The ordered observations are again denoted by $X_{(1)} \leq \dots \leq X_{(n)}$ ($X_{(1)}^* \leq \dots \leq X_{(n)}^*$). From Theorem 3.2 the following can be obtained.

Theorem 3.3. (Pledger and Proschan, 1971). Let $(-\log \bar{F}_1(t), \dots, -\log \bar{F}_n(t)) \geq^m (-\log \bar{F}_1^*(t), \dots, -\log \bar{F}_n^*(t))$ for each $t \geq 0$. Then $X_{(1)} =^{st} X_{(1)}^*$ and $X_{(k)} \geq^{st} X_{(k)}^*$ for $k = 2, \dots, n$.

Note that by interchanging F_i and \bar{F}_i and $P(X_k > t)$ and $P(X_{(n-k+1)} \leq t)$, we can derive the dual of Theorem 3.3: Let $(-\log F_1(t), \dots, -\log F_n(t)) \geq^m (-\log F_1^*(t), \dots, -\log F_n^*(t))$ for each $t \geq 0$. Then $X_{(k)} \leq^{st} X_{(k)}^*$ for $k = 1, \dots, n-1$ and $X_{(n)} =^{st} X_{(n)}^*$.

If $\bar{F}_i^*(t), i = 1, \dots, n$, are all equal to the geometric mean of the $\bar{F}_1(t), \dots, \bar{F}_n(t)$, then $X_{(k)} \geq^{st} X_{(k)}^*$ for $k = 1, \dots, n$; in particular, $X_{(1)} =^{st} X_{(1)}^*$. Likewise, if $F_1^*(t) = \dots = F_n^*(t) = [\prod_{i=1}^n F_i(t)]^{1/n}$ for $t \geq 0$, then $X_{(k)} \leq^{st} X_{(k)}^*$ for $k = 1, 2, \dots, n$; in particular, $X_{(n)} =^{st} X_{(n)}^*$.

In keeping with (3.1) of hazard in the non time-dependent case, we define the hazard function $R(t)$ corresponding to survival probability $\bar{F}(t)$ in the time-dependent case by

$$R(t) = -\log \bar{F}(t) \text{ for } t \geq 0. \quad (3.4)$$

We say hazards are proportional if hazard $R_i(t)$ may be expressed as

$$R_i(t) = \lambda_i R(t) \text{ for } t \geq 0, \lambda_i > 0, i = 1, \dots, n, \quad (3.5)$$

where $R(t)$ is a hazard function.

The concept of proportional hazard functions is a very useful one in reliability theory. Assume that the heterogeneous distributions have proportional hazards $R_i(t) = \lambda_i R(t)$ and $R_i^*(t) = \lambda_i^* R(t), i = 1, 2, \dots, n$, where $R(t)$ is a hazard function and that $\underline{\lambda} \geq^m \underline{\lambda}^*$. Then one can easily see that $\underline{\lambda} \geq^m \underline{\lambda}^*$ implies

$$(-\log \bar{F}_1(t), \dots, -\log \bar{F}_n(t)) \geq^m (-\log \bar{F}_1^*(t), \dots, -\log \bar{F}_n^*(t))$$

for each $t \geq 0$. Therefore, by Theorem 3.3, we have $X_{(1)} =^{st} X_{(1)}^*$ and $X_{(k)} \geq^{st} X_{(k)}^*, k = 2, \dots, n$. We state this result in the following theorem.

Theorem 3.4 (Pledger and Proschan, 1971). Let $F_1, \dots, F_n; F_1^*, \dots, F_n^*$ have proportional hazard functions with $\lambda_1, \dots, \lambda_n; \lambda_1^*, \dots, \lambda_n^*$ as constants of proportionality. Let $\underline{\lambda} \geq^m \underline{\lambda}^*$. Then $X_{(1)} =^{st} X_{(1)}^*$ and $X_{(k)} \geq^{st} X_{(k)}^*, k = 2, \dots, n$.

Another interesting stochastic comparison can be obtained when survival probability is logarithmically convex in the parameter λ . Consider survival probability $\bar{F}(t, \lambda) = \bar{G}(\lambda t)$, where G has DFR and λ occurs as a scale factor. It is a well known fact that a DFR survival probability is log convex (See Barlow and Proschan, 1981, Chapter 3). Some examples of DFR survival probability are Weibull and gamma when shape parameters are ≤ 1 . Also a mixture of exponential distributions has DFR; see Barlow and Proschan (1981), Chapter 4.

Theorem 3.5. (Pledger and Proschan, 1971). For $t \geq 0$ let $(\bar{F}(t, \lambda_i))$ be differentiable, monotone, and log convex in $\lambda_i \geq 0, i = 1, 2, \dots, n$. If $\underline{\lambda} \geq^m \underline{\lambda}^*$, then $X_{(k)} \geq^{st} (\leq^{st}) X_{(k)}^*$ for $k = 1, \dots, n$.

Furthermore, if $F(t, \lambda_i)(\bar{F}(t, \lambda_i))$ is differentiable and log concave in $\lambda_i \geq 0, i = 1, \dots, n$, and $\underline{\lambda} \geq^m \underline{\lambda}^*$, then $X_{(n)} \geq^{st} X_{(n)}^*(X_{(1)} \leq^{st} X_{(1)}^*)$.

The above comparisons for the largest and smallest order statistics also hold for parallel and series systems.

Next we discuss comparisons of spacings arising from a single set of order statistics as well as the spacings arising from heterogeneous distributions. We assume the underlying heterogeneous distributions have proportional hazards $\lambda_1 R(t), \dots, \lambda_n R(t)$.

Let $D_1 = X_{(1)}, D_2 = X_{(2)} - X_{(1)}, \dots, D_n = X_{(n)} - X_{(n-1)}$ denote the spacings between order statistics of independent observations, one from each of n heterogeneous distributions. We may recall that when the observations come from a single underlying exponential distribution F , then the normalized spacings $nD_1, (n-1)D_2, \dots, D_n$ are independently distributed according to the same exponential distribution F and are stochastically alike, i.e., $nD_1 =^{st} (n-1)D_2 =^{st} \dots =^{st} D_n$. More general results have been obtained by Pledger and Proschan (1971) when the underlying distributions have concave proportional hazards. They show the normalized spacings increase stochastically.

Theorem 3.6. Let $\bar{F}_i(t) = e^{-\lambda_i R(t)}, \lambda_i > 0$ for $i = 1, \dots, n$, where $R(t)$ is concave and differentiable. Then $nD_1 \leq^{st} (n-1)D_2 \leq^{st} \dots \leq^{st} D_n$.

Define $D_i^*, i = 1, \dots, n$ to be the spacing arising from $\{F_1^*, \dots, F_n^*\}$, the second set of exponential distributions. Pledger and Proschan (1971) obtain the following comparisons of two sets of spacings.

Theorem 3.7. Let $\bar{F}_i(t) = e^{-\lambda_i t}$, and $\bar{F}_i^*(t) = e^{-\bar{\lambda} t}$, for $i = 1, 2, \dots, n$, where $\bar{\lambda}$ is the arithmetic mean of the $\lambda_1, \dots, \lambda_n$. Then $D_1 =^{st} D_1^*$ and $D_i \geq^{st} D_i^*$ for $i = 2, \dots, n$.

Note that the mean $\bar{\lambda}$ is used for the comparison. A natural question to ask is: Is the same conclusion possible using a vector $\underline{\lambda}^*$, where $\underline{\lambda} \geq^m \underline{\lambda}^*$? The conclusion need not hold when $\underline{\lambda} \geq^m \underline{\lambda}^*$. A counterexample is provided in their paper.

An application of the results of this section to a reliability problem is as follows. A k -out-of- n system is to be designed from supposedly like units. However, due to random fluctuations in the units, the individual unit reliabilities actually vary. From a knowledge of the average reliability of the units, we wish to predict conservatively the reliability of the system of unlike units. The

theorems in this section describe conditions under which we may obtain such a conservative prediction.

4. STOCHASTIC COMPARISONS OF VECTORS OF ORDER STATISTICS.

The results discussed so far present stochastic comparisons between an individual order statistic from heterogeneous populations and the corresponding statistic from a homogeneous population.

In this section we review stochastic comparisons of pairs of vectors of order statistics from heterogeneous populations. This stochastic vector comparison yields as special cases the Pledger-Porschans theorems described above and additional results stochastically comparing partial or complete sums of order statistics.

Definition 4.1. A function f defined on R^n is said to be increasing if it is increasing in each argument.

Definition 4.2. The random vector $\underline{X} = (X_1, \dots, X_n)$ is said to be stochastically smaller than the random vector $\underline{X}^* = (X_1^*, \dots, X_n^*)$, (denoted by $\underline{X} \leq^{st} \underline{X}^*$) if $f(\underline{X}) \leq^{st} f(\underline{X}^*)$ for every real-valued, Borel measurable, increasing function f defined on R^n .

It is well known that $\underline{X} \leq^{st} \underline{X}^*$ if and only if $P(\underline{X} \in B^n) \leq P(\underline{X}^* \in B)$ for every upper open subset B of R^n . (A set $B \subset R^n$ is said to be an upper set if $\underline{a} \in B$ and $\underline{a} \leq \underline{b}$ implies $\underline{b} \in B$.)

Sufficient conditions for the stochastic comparison of two random vectors are given by Shanthikumar (1987). His conditions are weaker than those of Veinott (1965), and of Arjas and Lehtonen (1978).

Given a vector $\underline{X} = (X_1, \dots, X_n)$, let $X_{(1)} \leq \dots \leq X_{(n)}$ denote an increasing rearrangement of the coordinates. We denote the vector $(X_{(1)}, \dots, X_{(n)})$ by $\underline{X}_{(1)}$. Proschan and Sethuraman (1976) investigate stochastic comparisons of various functions of $\underline{X}_{(1)} = (X_{(1)}, \dots, X_{(n)})$ with similar functions $\underline{X}_{(1)}^* = (X_{(1)}^*, \dots, X_{(n)}^*)$ under suitable conditions on $F_1, F_2, \dots, F_n; F_1^*, F_2^*, \dots, F_n^*$. We state the main theorem.

Theorem 4.1. (Proschan and Sethuraman, 1976). Let $F_1, \dots, F_n; F_1^*, \dots, F_n^*$ have proportional hazard functions with $\lambda_1, \dots, \lambda_n; \lambda_1^*, \dots, \lambda_n^*$ as the constants of proportionality. Let $\lambda \geq^m \lambda^*$. Then $\underline{X}_{(1)} \geq^{st} \underline{X}_{(1)}^*$. A special case involving exponential distributions can be readily obtained.

Theorem 4.2. (Proschan and Sethuraman, 1976). Let $Y_1, \dots, Y_n (Y_1^*, \dots, Y_n^*)$ be independent exponential random variables with parameters $\lambda_1, \dots, \lambda_n (\lambda_1^*, \dots, \lambda_n^*)$, respectively. Let $\underline{\lambda} \geq^m \underline{\lambda}^*$. Then $\underline{Y}_{(r)} \geq^{st} \underline{Y}_{(r)}^*$.

Theorem 4.1 has important applications in reliability and life testing which will be discussed later. An interesting and useful special case of Theorem 4.1 is given as follows : Since $\sum_{r \in I} X_{(r)}$ is an increasing function of $\underline{X}_{(r)}$ for each subset I of $\{1, 2, \dots, n\}$ under the conditions of Theorem 4.1 we have $\sum_{r \in I} X_{(r)} \geq^{st} \sum_{r \in I} X_{(r)}^*$. Thus $\sum_{r=1}^k X_{(r)} \geq^{st} \sum_{r=1}^k X_{(r)}^*$ for $r = 1, 2, \dots, k$; in particular $\sum_{r=1}^n X_{(r)} \geq^{st} \sum_{r=1}^n X_{(r)}^*$.

Note that Theorem 3.4 is an immediate consequence of Theorem 4.1, since $X_{(r)}$ is an increasing function of $\underline{X}_{(r)}$ for $r = 1, \dots, n$, and also note that in Theorem 3.4 the order statistics are stochastically compared one at a time.

It can be checked easily from the distribution functions of the X and X^* that if $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and $\lambda_1^* \geq \dots \geq \lambda_n^*$, then

$$X_1 \leq^{st} \dots \leq^{st} X_n$$

and

$$X_1^* \leq^{st} \dots \leq^{st} X_n^*$$

The stochastic ordering above has been achieved by ordering the parameters of the distributions. One may ask whether $\underline{\lambda} \geq^m \underline{\lambda}^*$ implies $\sum_{i=k}^n X_i \geq^{st} \sum_{i=k}^n X_i^*$. An affirmative answer is given by Proschan and Sethuraman (1976) in the following.

Theorem 4.3 Let $\lambda_1 \geq \dots \geq \lambda_n (\lambda_1^* \geq \dots \geq \lambda_n^*)$. Let $X_1, \dots, X_n (X_1^*, \dots, X_n^*)$ be independent random variables with proportional hazard functions and with constants of proportionality $\lambda_1, \dots, \lambda_n (\lambda_1^*, \dots, \lambda_n^*)$, respectively. Let $\underline{\lambda} \geq^m \underline{\lambda}^*$. Then for each $k, 1 \leq k \leq n$,

$$\sum_{i=k}^n X_i \geq^{st} \sum_{i=k}^n X_i^*.$$

Application 4.1. An important application of Theorem 4.1 concerns the robustness of standard estimators of the failure rate of an exponential distribution when observations are actually from heterogeneous distributions. Let Y_1, \dots, Y_n be i.i.d. exponential random variables with failure

rate λ . Under censored sampling in which observations cease at the r th failure (i.e., Type-II censoring), the standard estimator (UMVUE) $\hat{\lambda}$ takes the form

$$\hat{\lambda}(\underline{Y}) = r / \left[\sum_1^r Y_{(i)} + (n - r) Y_{(r)} \right], \quad (4.1)$$

where $Y_{(i)}$ is the i th order statistic $i = 1, \dots, n$.

Suppose now that the observations actually come from different exponential distributions, with average failure rate λ . To be specific, let Y_1^*, \dots, Y_n^* be independent exponential random variables with failure rates $\lambda_1^*, \dots, \lambda_n^*$, respectively, and let $\lambda = \sum_{i=1}^n \lambda_i^*/n$. Note that $\hat{\lambda}(\underline{Y})$ in (4.1) is a decreasing function of $\underline{Y}_{(r)}$. Thus it follows from Theorem 4.2 that

$$\hat{\lambda}(\underline{Y}) \geq^{st} \hat{\lambda}(\underline{Y}^*). \quad (4.2)$$

The implication of (4.2) is that the estimate $\hat{\lambda}$ in (4.1) tends to underestimate the average failure rate in the presence of heterogeneity. Proschan and Sethuraman (1976) note that Theorem 4.2 actually gives a more refined conclusion: The greater the degree of heterogeneity (as reflected by majorization) among $\lambda_1^*, \dots, \lambda_n^*$ satisfying $\lambda = \sum_{i=1}^n \lambda_i^*/n$, the greater the underestimation of λ .

Barlow and Proschan (1967) have listed estimates for λ under various sampling schemes. These estimates are of the form:

$$\hat{\lambda}(\underline{Y}) = \frac{\text{number of failures observed}}{\text{total time on test}}. \quad (4.3)$$

It can be seen (See Barlow and Proschan, 1967) that in general the estimate in (4.3) is a decreasing function of the order statistics. Thus, as in the above case of censored sampling, heterogeneity of the exponential distributions leads to underestimation of the average failure rate when using the estimate in (4.3).

The following direct applications of Theorem 4.1 are discussed by Proschan and Sethuraman (1976).

Application 4.2. Let X_1, X_2, \dots be i.i.d random variables having a Weibull distribution with shape parameter $\alpha > 0$; i.e., $P(X_1 > x) = e^{-x^\alpha}$, $x > 0$. Let $\underline{u} = (u_1, \dots, u_n)$ and $\underline{u}^* = (u_1^*, \dots, u_n^*)$ be vectors such that $(u_1^{-\alpha}, \dots, u_n^{-\alpha}) \geq^{st} ((u_1^*)^{-\alpha}, \dots, (u_n^*)^{-\alpha})$. Then

$$\sum_{i=1}^n u_i X_i \geq^{st} \sum_{i=1}^n u_i^* X_i. \quad (4.4)$$

From the Weibull distribution it is easy to see that $u_1 X_1, \dots, u_n X_n$ have proportional hazard functions with constants of proportionality $u_1^{-\alpha}, \dots, u_n^{-\alpha}$, respectively. Thus (4.4) is immediate from Theorem 4.1.

Application 4.3. Let X_p denote the binomial random variable with $P(X_p = 1) = 1 - P(X_p = 0) = p$. Let X_{p_1}, \dots, X_{p_n} ($X_{p_1^*}, \dots, X_{p_n^*}$) be mutually independent and let $(-\log p_1, \dots, -\log p_n) \geq^m (-\log p_1^*, \dots, -\log p_n^*)$. Then $\sum_{i=1}^n X_{p_i} \geq^{st} \sum_{i=1}^n X_{p_i^*}$.

We note that Theorem 3.2 (Pledger and Proschan) concerning the reliability of k -out-of- n systems and order statistics can be obtained as a consequence of Application 4.3, since $h_k(\underline{p}) = P(\sum_{i=1}^n X_{p_i} \geq k)$.

The usual definition of stochastic comparison of random vectors has been extended by Pledger and Proschan (1973) to stochastic comparison of random processes. We call stochastic process $\{X(t), t \geq 0\}$ stochastically larger than stochastic process $\{Y(t), t \geq 0\}$ (written $\{X(t), t \geq 0\} \geq^{st} \{Y(t), t \geq 0\}$) if $(X(t_1), \dots, X(t_n)) \geq^{st} (Y(t_1), \dots, Y(t_n))$ for every choice of $0 \leq t_1 < t_2 < \dots < t_n, n = 1, 2, \dots$. This extended comparison permits one to obtain bounds not just on a few parameters, but simultaneously on an uncountably infinite class of functionals of the stochastic process. See Pledger and Proschan (1973) for applications to reliability problems, yielding stochastic comparisons for systems of independently operating machines assuming exponential failure and exponential repair.

REFERENCES.

- Arjas, E. and Lehtonen, T. (1978). Approximating many server queues by means of single server queues. *Math. Operat. Res.* **3**, 205-223.
- Barlow, R. E., Marshall, A. W. and Proschan, F. (1969). Some inequalities for starshaped and convex functions. *Pacific J. Math.* **29** 19-42.
- Barlow, R. E. and Proschan, F. (1966a). Inequalities for linear combinations of order statistics from restricted families. *Ann. Math. Statist.* **37** 1574-1592.
- Barlow, R. E. and Proschan, F. (1966b). Tolerance and confidence limits for classes of distributions based on failure rate. *Ann. Math. Statist.* **37** 1593-1601.
- Barlow, R. E. and Proschan, F. (1967). Exponential life test procedures when the distribution has monotone failure rate. *J. Amer. Statist. Assoc.* **62** 548-560.
- Barlow, R. E. and Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing: Probability Models*. To Begin With, Silver Spring, MD.
- Bruckner, A. M. and Ostrow, E. (1962). Some function classes related to the class of convex functions. *Pacific J. Math.* **12** 1203-1215.
- Chan, W., Proschan, F., and Sethuraman, J. (1987). Schur-Ostrowski theorems for functionals in $L_1(0, 1)$. *SIAM J. Math. Analysis.* **18** 566-575.
- Chan, W., Proschan, F., and Sethuraman, J. (1983). Convex-ordering among distributions with applications to reliability and mathematical statistics. FSU Technical Report No. M-661.
- David, H. A. (1970). *Order Statistics*. John Wiley and Sons, Inc.
- David, H. A. (1986). Inequalities for ordered sums. *Ann. Inst. Statist. Math.* **38** 551-555.
- Epstein, B. (1960a). Statistical life test acceptance procedures. *Technometrics*. **2** 435-446.
- Epstein, B. (1960b). Estimation from life test data. *Technometric*. **2** 447-454.
- Epstein, B. and Sobel, M. (1953). Life testing. *J. Amer. Statist. Assoc.* **48** 486-502.
- Groeneveld, R. A. (1982). Best bounds for order statistics and their expectations in range and mean units with applications. *Comm. Statist. - Theory Meth.* **11** 1809-1815.
- Karlin, S. (1968). *Total Positivity*. Stanford University Press.
- Marshall, A. W. and Olkin, I. (1979). *Inequalities: theory of Majorization and Its Applications*. Academic Press.

- Marshall, A. W., Olkin, I. and Proschan, F. (1967). Monotonicity of ratios of means and other applications of majorization. *Inequalities*. Ed. O. Shisha, Academic Press Inc. 177-190.
- Ostrowski, A. (1952). Sur quelques applications des fonctions convexes et concaves au sens de I. Schur. *J. Math. Pures Appl.* **31** 253-292.
- Pledger, G. and Proschan, F. (1971). Comparisons of order statistics and of spacings from heterogeneous distributions. In *Optimizing Methods in statistics*. (J. S. Rustagi, Ed.) Academic Press, New York.
- Pledger, G. and Proschan, F. (1973). Stochastic comparisons of random processes, with applications in reliability. *J. Appl. Prob.* **10** 572-585.
- Proschan, F. and Sethuraman, J. (1976). Stochastic comparisons of order statistics from heterogeneous populations, with applications in reliability. *J. Mult. Anal.* **6** 608-616.
- Pyke, R. (1965). Spacings. *J. R. Statist. Soc. B* **27** 395-436. Discussion: 437-449.
- Pyke, R. (1970). Spacings revisited, Technical Report No. 25, Department of Mathematics, University of Washington.
- Ryff, J. V. (1963). On the representation of doubly stochastic operators. *Pacific J. Math.* **13** 1379-1386.
- Sarhan, A. E. and Greenberg, B. G. (Eds.) (1962). *Contributions to Order Statistics*. John Wiley and Sons, Inc.
- Schur, I. (1923). Über ein Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzber. Berl. Math. Ges.* **22** 9-20.
- Sen, P. K. (1970). A note on order statistics for heterogeneous distributions. *Ann. Math. Statist.* **41** 2137-2139.
- Shanthikumar, J. G. (1987). On stochastic comparison of random vectors. *J. Appl. Prob.* **24** 123-136.
- Veinott, A. F. (1965). Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. *Operat. Res.* **13** 761-778.
- Van Zwet, W. R. (1964). *Convex Transformations of Random Variables*. Mathematisch Centrum.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER <i>ARO 23699.16-MA</i>	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Stochastic Comparisons of Order Statistics, with Applications in Reliability		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER FSU Technical Report M-773
7. AUTHOR(s) Jee Soo Kim, Frank Proschan and Jayaram Sethuraman		8. CONTRACT OR GRANT NUMBER(s) DAAL03-86-K-0094
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University Department of Statistics Tallahassee, FL 32306-3033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE November, 1987
		13. NUMBER OF PAGES 18
		14. SECURITY CLASS. (of this report) UNCLASSIFIED
		15. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) N/A		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Stochastic comparison, order statistics, reliability, life testing, majorization, Schur function, heterogeneous distributions, k-out-of-n system, proportional hazard functions.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This is an invited paper to appear in the special issue of <u>Communications in Statistics: Theory and Methods</u> devoted to "Order Statistics and Applications". This paper reviews recent developments in stochastic comparisons of order statistics.		

END
DATE

FILMED
MARCH
1988

DTIC